Hearing Report: Professor Antonio Casilli Spain Expert Committee on Democracy at Work, 2025

May 23, 2025

The hearing report draws on extensive evidence provided to the Commission by Antonio Casilli, Professor of Sociology at the Institute Polytechnique de Paris, in his testimony on "The Crisis of Informality and the Global Value Chain" for the Expert Committee on Democracy at Work, established by the Ministry of Labour of the Government of Spain (Friday, May 23rd, 2025).

Professor Antonio Casilli's evidence to the Commission addressed a frequently under-examined aspect which could be seen as a preliminary stage to Algorithmic Management: the externalized labour that remains largely concealed to the public eye yet plays a crucial role in artificial intelligence design and development. Drawing from extensive research programme conducted with his DiPLab team (Digital Platform Labor), the presentation focused specifically on data work as an essential but undervalued component of AI production systems.

Defining Data Work

Analysis was structured around the concept of data work existing on a continuum with other forms of platform labor rather than as discrete categories. At one end of this spectrum lies gig work, exemplified by location-based, on-demand platforms such as Uber and Glovo. Moving along the continuum, we find remote data work, which operates through either platform-based systems or business process outsourcing arrangements and is characterized by consistent underpayment relative to the value generated. The far end of the continuum encompasses unpaid data activities (such as customers' completing only

Captchas), which lies beyond the scope of the Commission's enquiry.

Our focus, Professor Casilli suggested, should be on workers who engage in the fundamental processes that enable AI systems to function: they are frequently responsible for preparing datasets, training algorithms, validating model outputs, and evaluating system performance. Professor Casilli identified four primary categories of tasks performed by data workers within AI production systems. This includes labelling activities, where they assign categorical tags to various forms of data to enable machine learning algorithms to recognize patterns. Annotation represents another crucial function, involving more detailed descriptive work that provides context and meaning to raw data inputs. Content moderation constitutes a third major area of responsibility, requiring workers to review and assess material for appropriateness, accuracy, or policy compliance. Finally, Professor Casilli identified content filtering as a key task, involving the systematic removal or flagging of problematic material before it can influence AI system behaviour.

Despite the critical nature of their contributions to the AI value chain, this work typically remains invisible to end users and is systematically undervalued within the broader technology ecosystem. Disparities in compensation and working conditions are stark: workers in Kenya who received compensation as low as $1.34 per hour for their contributions to AI training processes are far from alone, with hidden AI labour involving workers across multiple countries including South Africa, the United Kingdom, Turkey, India, Canada, the United States, and the Philippines. These are only the countries mentioned by OpenAI about GPT-3 training, while DiPLab research has conducted research across several Spanish- and Portuguese-speaking countries in Latin America (Venezuela, Colombia, Brazil, Mexico, Argentina), French- and English-speaking countries in Africa (Madagascar, Kenya, Cameroon, Côte d'Ivoire, Uganda, Egypt), and of course South- and South-East Asia (India, Philippines, Bangladesh, Nepal, and of course China).

Data Workers and the Global North-South Divide

One of the key themes of the presentation was the geographic distribution of AI labour, revealing what he characterized as a fundamental North-South divide in the industry. Whilst artificial intelligence appears to be 'produced' in the Global North, where major technology companies maintain their headquarters and operate large-scale data centres,

Professor Casilli's research demonstrated that this appearance masks a more complex reality.

This was illustrated drawing on field research findings from Madagascar, a country heavily involved in outsourced data labour specifically serving French firms. In addition to poor infrastructure, internet access, and shared or inadequate facilities, the research documented extremely low monetary compensation or, in many cases, complete absence of monetary payment structures, with workers paid in food items such as sugar, rice, and beans instead of cash for their data labour.

Mapping the Primary Data Labour Flow Patterns

According to Professor Casilli's analysis, data production remains concentrated in the Global South while AI companies maintain their corporate headquarters and public presence in the Global North, including major corporations such as Meta and Microsoft. This geographic separation enables data to move globally while remaining largely unrecognized in official corporate narratives about AI development processes.

This structural arrangement, as Professor Casilli described it, can be broken down into four distinct flow patterns that characterize the contemporary AI production landscape.

The first flow pattern involves movement from South and Southeast Asia to the United States. Professor Casilli noted that this pattern builds upon traditional outsourcing relationships, with established hubs in countries including India and the Philippines serving large U.S. technology companies. A second pattern involves internal Chinese data flows, where labour moves from what he characterized as Tier-3 cities—typically low-income, densely populated urban areas—to coastal AI development hubs including Shenzhen and Shanghai. Professor Casilli noted that this pattern also incorporates overseas Chinese-speaking communities, particularly in Malaysia and Singapore.

The third pattern documented by Professor Casilli involves Latin American countries, including Argentina, Venezuela, Colombia, and Brazil, producing data for both U.S. and European firms, with particular emphasis on Spanish companies, reflecting both linguistic connections and geographic ties. A fourth pattern, finally, centres on African countries serving global clients through multiple linguistic and colonial connections. According to Professor Casilli's findings, English-speaking African countries such as Kenya and

Nigeria primarily serve U.S. firms, while French-speaking countries including Tunisia, Côte d'Ivoire and Madagascar focus on French corporate clients. Professor Casilli noted that African workers also increasingly provide data labour services to Chinese firms, including specifically in the field of facial recognition development.

Policy Responses

In concluding his presentation, Professor Casilli outlined three strategic responses to address what he characterized as the data labour crisis: downstream accountability measures, citing initiatives such as Fairwork and the Partnership on AI that aim to push companies and platforms toward meeting established ethical standards; upstream pressure as a complementary approach, emphasizing worker-led efforts including union organizing, legal action, and grassroots activism that challenge existing power structures from below; and alternative governance models, particularly platform cooperatives and community ownership structures. In terms of concrete examples, Professor Casilli discussed the EU Corporate Sustainability Due Diligence Directive, which seeks to establish legal accountability for companies regarding their supply chains, noting its incomplete implementation across member states and Spain's potential leadership role, and the ILO Labour Disclosure Initiative, which would mandate transparency requirements including disclosure of AI worker numbers, task descriptions, and working conditions.

Professor Casilli emphasized that none of these approaches alone would prove sufficient, given the complex and systemic nature of the underlying problems and intersectional challenges, including traditional labour rights and employment precarity, persistent gender inequality, systemic lack of skill recognition, and barriers related to migration status and citizenship.

References

Casilli, A. A. (2017) 'Digital Labor Studies Go Global: Toward a Digital Decolonial Turn,' *International Journal of Communication,* 11, Special Section 'Global Digital Culture,' pp. 3934–3954.

Casilli, A. A. (2025) *Waiting for Robots. The Hired Hands of Automation.* Chicago:

University of Chicago Press.

Casilli, A. A., Tubaro, P., Cornet, M., Le Ludec, C., Torres-Cierpe, J., et al. (2025) *Global Inequalities in the Production of Artificial Intelligence: A Four-Country Study on Data Work.* In J. Qiu, S. Yeo and R. Maxwell (eds.), *The Handbook of Digital Labor.* Wiley Blackwell, pp. 219–232.

Miceli, M., Tubaro, P., Casilli, A. A., Le Bonniec, T., Wagner, C. S., et al. (2024) *Who Trains the Data for European Artificial Intelligence?* European Parliament; The Left, pp. 1–40. ⟨hal-04662589⟩

Muldoon, J., Graham, M. and Callum, C. (2024) *Feeding the Machine: The Hidden Human Labour Powering AI.* London: Bloomsbury.

Tubaro, P., Casilli, A. A., Cornet, M., Le Ludec, C. and Torres-Cierpe, J. (2025) 'Where Does AI Come From? A Global Case Study Across Europe, Africa, and Latin America,' *New Political Economy,* Online First, pp. 1–14.